

### REMARKS

In response to the Office Action mailed June 9, 2009 (hereinafter "the Office Action"), the Assignee respectfully requests reconsideration. Claims 1-15 and 17-27 were previously pending in the application. In this amendment, no claims are amended, added, nor cancelled. As a result, claims 1-15 and 17-27 are pending for examination with claims 1, 6, 11, 17, 21, and 24 being independent.

#### Claim Rejections Under 35 U.S.C. § 103: Claims 1-15

Beginning at page 3, the Office Action rejects claims 1-16 under 35 U.S.C. § 103(a) as purportedly being unpatentable over Neti (U.S. Patent No. 5,953,701) in view of Naito (U.S. Patent No. 5,983,178) in further view of Kanevsky (U.S. Patent No. 6,529,902). Please note, claim 16 was cancelled in the Assignee's prior response. The Assignee respectfully traverses the rejections of claims 1-15, of which claims 1, 6, and 11 are independent claims, for the reasons set forth below.

#### Claim 1

Claim 1 is directed to a computer readable medium product encoded with instructions for generating a speech recognition model from female, male and gender-independent phoneme models. Claim 1 recites, *inter alia* and in combination with other elements of the claim, "determining a difference between each female phoneme model and each corresponding male phoneme model; creating a gender-independent phoneme model when the difference between the compared female phoneme model and the corresponding male phoneme model is less than a predetermined value; and adding, based on a criterion, one of the gender-independent phoneme model, or both the female phoneme model and the corresponding male phoneme model to the speech recognition model." Neti, Naito, and Kanevsky have been carefully reviewed, particularly those sections cited in the Office Action. None of these references, considered alone or together, at least disclose these elements of claim 1, and therefore claim 1 patentably distinguishes over Neti, Naito, and Kanevsky. Reasons supporting this position are set forth below.

The basis relying on Neti set forth in the Office Action for rejecting claim 1 was persuasively overcome in Assignee's amendment filed November 23, 2009. In particular, Neti fails

at least to disclose “determining a difference between each female phoneme model and each corresponding male phoneme model.” The Office Action now cites the Abstract, Col. 3, lines 37-49, and Col. 4 lines 10-29 as disclosing this element of claim 1. The sections have been reviewed carefully, and a disclosure of this element cannot be found. At best, the cited sections relate to constructing a hybrid gender-dependent model which is guided by evaluating phonetic context dependence. The cited sections disclose that gender-dependent and context-dependent decision trees are used in the process. There is, however, no disclosure anywhere in Neti of “determining a difference between each female phoneme model and each corresponding male phoneme model” as recited in claim 1 and construed in the context of the claim as a whole. The arguments set forth by the Assignee in the response filed November 23, 2009 in connection with this same issue still apply. Neither Naito nor Kanevsky cure the deficiencies of Neti, and therefore at least this element of claim 1 is not disclosed by Neti, Naito, nor Kanevsky.

The Office Action acknowledges that Neti fails to disclose “creating a gender-independent phoneme model when the difference between the compared female phoneme model and the corresponding male phoneme model is less than a predetermined value,” and purports that Naito discloses this element of the claim. The Office Action copies a cited section of Naito (Col. 3, line 56 – Col. 4, line 12), which fails to disclose “creating a gender-independent phoneme model when the difference between the compared female phoneme model and the corresponding male phoneme model is less than a predetermined value.” The copied section does not even disclose creating a model based on the comparison of two other models. The other cited section of Naito (Col. 15, line 54 – Col. 16, line 25) pertains to phoneme verification during speech recognition (which is not model building) in which a “verification score” and a “phoneme recognition score” is returned. Again, there is nothing in the cited section disclosing “creating a gender-independent phoneme model when the difference between the compared female phoneme model and the corresponding male phoneme model is less than a predetermined value.” Therefore, claim 1 patentably distinguishes over Neti and Naito.

The Office Action mischaracterizes Naito at Col. 15, line 54 – Col. 16, line 25 in stating that it “teaches the recognition of phoneme dependent data which verifies whether data is independent of dependent, for example whether incoming data is within a range of a model or not.” (*sic*) There is

nothing in this section about dependent or independent phonemes. There is also nothing in the section about “whether incoming data is within a range of a model or not.” The immediately preceding section (Col. 15, lines 45-53) teaches that a “cluster model to be used in the phoneme verifier” has already been designated.

Kanevsky fails to cure the deficiencies of Neti and Naito. Kanevsky is relied on (p. 9) for purportedly disclosing “a threshold value or an upper limit for the total number of phoneme models.” However, Kanevsky fails to disclose “creating a gender-independent phoneme model when the difference between the compared female phoneme model and the corresponding male phoneme model is less than a predetermined value; and adding, based on a criterion, one of the gender-independent phoneme model, or both the female phoneme model and the corresponding male phoneme model to the speech recognition model,” as recited in claim 1. Therefore claim 1 is patentable over Neti, Naito, and Kanevsky.

In view of the foregoing, Neti, Naito, and Kanevsky fail at least to disclose every element of claim 1. Therefore, claim 1 patentably distinguishes over Neti, Naito, and Kanevsky. Accordingly, withdrawal of the rejection of claim 1 under 35 U.S.C. § 103(a) is respectfully requested.

Claims 2-5 depend from claim 1, and are therefore allowable for at least the same reasons.

#### Claim 6

Claim 6 encompasses a system for generating speech recognition models and recites, “determine a difference between each first phoneme model and each corresponding second phoneme model; create an independent phoneme model when the difference between the compared each first phoneme model and each corresponding second phoneme model is less than a predetermined value; and add, based upon a criterion, one of the independent phoneme model, or both the first phoneme model and the corresponding second phoneme model to the speech recognition model.” For reasons that should be clear from the discussion of claim 1 above in connection with Neti, Naito, and Kanevsky, these cited references fail to teach or suggest every element of claim 6. Therefore, claim 6 patentably distinguishes over Neti, Naito, and Kanevsky. Accordingly, withdrawal of the rejection of claim 6 under 35 U.S.C. § 103(a) is respectfully requested.

Claims 7-10 depend from claim 6 and are therefore allowable for at least the same reasons.

#### Claim 11

Claim 11 encompasses a computer program product and recites, “determine a difference between each first phoneme model and each second phoneme model; create an independent phoneme model when the difference between the each first phoneme model and the each corresponding second phoneme model is less than a predetermined value; and add, based on a criterion, one of the independent phoneme model, or both the first phoneme model and the corresponding second phoneme model to the speech recognition model.” For reasons that should be clear from the discussion of claim 1 above in connection with Neti, Naito, and Kanevsky, these cited references fail to teach or suggest every element of claim 11. Therefore, claim 11 patentably distinguishes over Neti, Naito, and Kanevsky. Accordingly, withdrawal of the rejection of claim 11 under 35 U.S.C. § 103(a) is respectfully requested.

Claims 12-15 depend from claim 11 and are therefore allowable for at least the same reasons.

#### Claim Rejections Under 35 U.S.C. § 103: Claims 17-27

Beginning at page 17, the Office Action rejects claims 17-27 under 35 U.S.C. §103(a) as purportedly being unpatentable over Neti in view of Wark (U.S. Patent Publication No. 2003/0231775) and in further view of Naito. The cited references have been reviewed, and the Assignee respectfully traverses the rejections of claims 17-27.

#### Claim 17

Claim 17 is directed to a computer readable medium product encoded with instructions and recites, “computing best estimates that the current feature vector belongs to each one of the plurality of data classes; computing accumulated confidence values for each of the plurality of data classes that the current feature vector belongs to each one of the plurality of data classes, the confidence value for each data class based on the current best estimate and on previous confidence values for the each data class; weighing the class-dependent phoneme models based on the accumulated

confidence values; and recognizing the current feature vector based on the weighted class-dependent phoneme models.” At least these elements of claim 17 are not taught or suggested in Neti, Wark, or Naito, and therefore claim 17 patentably distinguishes over Neti, Wark, and Naito. Reasons supporting this position are set forth below.

The Office Action (p. 17) cites Neti at Col. 3, lines 50-67 and Col. 5, lines 9-21 as purportedly teaching “a first computing module configured to compute a current best estimates that the current feature vector belongs to one of the plurality of data classes.” (*sic*) This, however, is not what the claim recites. Claim 17 recites, “computing best estimates that the current feature vector belongs to each one of the plurality of data classes.” (emphasis added) Col. 3, lines 50-67 of Neti is directed to the construction of a context-dependent sub-phonetic model using binary questions and a decision-tree process. There is nothing in this section about computing best estimates that a current feature vector belongs to each one of a plurality of data classes. Col. 5, lines 9-21 discloses that some analyzed phonemes exhibit gender dependence as represented in Fig. 3 of Neti. Fig. 3 shows the percentages each of analyzed consonants, semivowels, vowels, and diphthongs that exhibit gender dependence. Again, there is nothing in this section about computing best estimates that a current feature vector belongs to each one of a plurality of data classes. Therefore, contrary to the assertion in the Office Action, Neti fails to disclose “computing best estimates that the current feature vector belongs to each one of the plurality of data classes.” Therefore, claim 17 patentably distinguishes over Neti.

The Office Action (p. 19) relies on Wark as purportedly disclosing “computing accumulated confidence values for each of the plurality of data classes that the current feature vector belongs to each one of the plurality of data classes, the confidence value for each data class based on the current best estimate and on previous confidence values for the each data class; weighing the class-dependent phoneme models based on the accumulated confidence values; and recognizing the current feature vector based on the weighted class-dependent phoneme models.” However, the Office Action does not particularly point out where each of these elements can be found in Wark. The sections of Wark ([0094], Fig. 4, [0129], [0130], and [0146]) generally alluded to in the Office Action disclose a process for analyzing a homogeneous audio segment by parsing the segment into audio frames and audio clips and using extracted clip feature vectors in order to classify or identify

an audio segment. There is no disclosure in the cited sections or elsewhere in Wark of “computing accumulated confidence values for each of the plurality of data classes that the current feature vector belongs to each one of the plurality of data classes, the confidence value for each data class based on the current best estimate and on previous confidence values for the each data class; weighing the class-dependent phoneme models based on the accumulated confidence values; and recognizing the current feature vector based on the weighted class-dependent phoneme models” as recited in claim 17 and construed in conjunction with the claim as a whole.

Further, Wark is directed to analyzing homogeneous audio segments as noted above. Wark defines a homogeneous audio segment as follows: “A homogeneous segment is a segment only containing samples from a source having constant acoustic characteristics, such as from a particular human speaker, a type of background noise, or a type of music.” (Wark, [0047]) All the sections cited in the Office Action pertain to analysis of a homogeneous audio segment, and have nothing to do with class-dependent phoneme models, to which claim 17 is directed. Therefore, Wark does not even disclose “recognizing the current feature vector based on the weighted class-dependent phoneme models.” Therefore, Neti and Wark fail to disclose every element of claim 17. Accordingly, claim 17 patentably distinguishes over Neti and Wark.

The Office Action (p. 20) relies on Naito for purportedly teaching “creating a class-independent/dependent phoneme model,” and repeats text from the rejection of claim 1. However, claim 17 does not recite “creating a class-independent/dependent phoneme model.” Claim 17 is directed to recognizing speech, not creating class-independent/dependent phoneme models. No relevance of Naito to claim 17 has been set forth in the Office Action, and Naito appears to have been included in error.

In view of the foregoing, claim 17 patentably distinguishes over Neti, Wark and Naito. Accordingly, withdrawal of the rejection of claim 17 under 35 U.S.C. § 103(a) is respectfully requested. Claims 18-20 depend from claim 17, and are therefore allowable for at least the same reasons.

### Claim 21

Claim 21 encompasses a system for recognizing speech data and recites, “a first computing module configured to compute current best estimates that the current feature vector belongs to each one of the plurality of data classes; a second computing module configured to compute accumulated confidence values for each of the plurality of data classes that the current feature vector belongs to each one of the plurality of data classes, the confidence value for each data class of the plurality of data classes based on the current best estimate for the data class and on previous confidence values for the data class, the previous confidence values associated with previous feature vectors of the audio stream; a weighing module configured to weigh the class-dependent phoneme models based on the accumulated confidence values; and a recognizing module configured to recognize the current feature vector based on the weighted class-dependent phoneme models.” For reasons that should be clear from the discussion of claim 17 above in connection with Neti, Wark, and Naito, these cited references also fail to teach or suggest every element of claim 21. Therefore, claim 21 patentably distinguishes over Neti, Wark, and Naito. Withdrawal of the rejection of claim 21 under 35 U.S.C. § 103(a) is respectfully requested.

Claims 22-23 depend from claim 21 and are therefore allowable for at least the same reasons.

### Claim 24

Claim 24 encompasses a computer program product and recites, “compute best estimates that the current feature vector belongs to each one of the plurality of data classes; compute accumulated confidence values for each of the plurality of data classes that the current feature vector belongs to each one of the plurality of data classes, the confidence value for each data class of the plurality of data classes based on the current best estimate for the data class and on previous confidence values for the data class, the previous confidence values associated with previous feature vectors of the audio stream; weigh the class-dependent phoneme models based on the accumulated confidence values; and recognize the current feature vector based on the weighted class-dependent phoneme models.” For reasons that should be clear from the discussion of claim 17 above in connection with Neti, Wark, and Naito, these cited references also fail to teach or suggest every

element of claim 24. Therefore, claim 24 patentably distinguishes over Neti, Wark, and Naito. Withdrawal of the rejection of claim 24 under 35 U.S.C. § 103(a) is respectfully requested.

Claims 25-27 depend from claim 24 and are therefore allowable for at least the same reasons.

#### General Comments on Dependent Claims

Because each of the dependent claims depends from a base claim that is believed to be in condition for allowance, Assignee believes that it is unnecessary at this time to argue the further distinguishing features of all of the dependent claims. However, Assignee does not necessarily concur with the interpretation of the dependent claims as set forth in the Office Action, nor does Assignee concur that the basis for the rejection of any of the dependent claims is proper. Therefore, Assignee reserves the right to specifically address in the future the further patentability of the dependent claims not specifically addressed herein.




**CONCLUSION**

In view of the above amendment, Assignee believes the pending application is in condition for allowance.

Assignee believes no fee is due with this response. However, if a fee is due, please charge our Deposit Account No. 23/2825 under Docket No. N0484.70762US00 from which the undersigned is authorized to draw.

Dated: September 9, 2010

Respectfully submitted,

By   
Marc S. Johannes  
Registration No.: 64,978  
WOLF, GREENFIELD & SACKS, P.C.  
600 Atlantic Avenue  
Boston, Massachusetts 02210-2206  
617.646.8000